

AD-A062 972

CALIFORNIA UNIV BERKELEY OPERATIONS RESEARCH CENTER
MATHEMATICAL MODELS FOR MAINTAINABILITY.(U)
SEP 78 R E BARLOW

F/G 14/4

UNCLASSIFIED

ARO-13179.7-M

DAAG29-76-G-0042
NI

OF /
AD
A062 972



END
DATE
FILMED
2-79

DDC

LEVEL II

AR0

13179.7-m

18

19

12

6

MATHEMATICAL MODELS FOR MAINTAINABILITY.

9

Final Report. 1 Oct 75 - 30 Sep 78,

10

Richard E./Barlow

11

Sep 78

12 25p.

THIS DOCUMENT IS BEST QUALITY PRACTICABLE.
THE COPY FURNISHED TO DDC CONTAINED A
SIGNIFICANT NUMBER OF PAGES WHICH DO NOT
REPRODUCE LEGIBLY.

U. S. Army Research Office

15

✓ DAAG29-76-G-0042

Operations Research Center
University of California, Berkeley

DDC
RECEIVED
JAN 9 1979
D

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED

79 01 02 051

270 750

mt

AD A062972

DDC FILE COPY

DISCLAIMER NOTICE

**THIS DOCUMENT IS BEST QUALITY
PRACTICABLE. THE COPY FURNISHED
TO DDC CONTAINED A SIGNIFICANT
NUMBER OF PAGES WHICH DO NOT
REPRODUCE LEGIBLY.**

THE FINDINGS IN THIS REPORT ARE NOT TO BE
CONSTRUED AS AN OFFICIAL DEPARTMENT OF
THE ARMY POSITION, UNLESS SO DESIGNATED
BY OTHER AUTHORIZED DOCUMENTS.

79 01 02 051

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) MATHEMATICAL MODELS FOR MAINTAINABILITY		5. TYPE OF REPORT & PERIOD COVERED Final Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Richard E. Barlow		8. CONTRACT OR GRANT NUMBER(s) DAAG29-76-G-0042
9. PERFORMING ORGANIZATION NAME AND ADDRESS Operations Research Center University of California Berkeley, California 94720		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS P-13179-M
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE September 1978
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		13. NUMBER OF PAGES 25
		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Maintainability Availability Uptimes Downtimes		<div style="border: 1px solid black; padding: 5px;"> DISTRIBUTION TO: GPO <input checked="" type="checkbox"/> White Section DOD <input type="checkbox"/> Encl Section UNANNOUNCED <input type="checkbox"/> JUSTIFICATION..... BY..... DISTRIBUTION/AVAILABILITY CODE DOC. ATAIL. ORD./OR SPECIAL <div style="display: flex; align-items: center;"> <div style="border: 1px solid black; padding: 5px; margin-right: 5px;">A</div> <div style="border: 1px solid black; padding: 5px;">23 E4.</div> </div> </div>
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) (SEE ABSTRACT)		

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ABSTRACT

A summary of work performed under ARO grant DAAG29-76-G-0042 is given. Invited talks as well as published reports are summarized.

MATHEMATICAL MODELS FOR MAINTAINABILITY

Period: 1 October 1975 - 30 September 1978

Our main goal on this project has been to develop theory and analytical techniques for modeling and analyzing maintained systems. In order to accomplish this goal we have, in conjunction with the Lawrence Berkeley Laboratory, analyzed maintenance records of the nuclear accelerator Superhilac with respect to system availability and reliability ([1], [2]). This has afforded us the opportunity to try out new techniques for analyzing system maintenance data and has also suggested problems for further research. Our analysis technique in [1] consisted (after first screening the data) of using time series methods to detect time dependencies followed by total time on test plots of system and subsystem uptimes and downtimes. Recommendations relative to improved maintenance policies were then made based on this information. In [2] a more detailed analysis of the Superhilac Radio Frequency Subsystem was made. Computer programs used in the maintenance data analysis are given there.

Recent graphical multivariate techniques for analyzing failure and repair data were described in an invited talk at the *Twenty-fourth Conference On The Design of Experiments in Army Research, Development and Testing*, at the Mathematics Research Center, Madison, Wisconsin, 4-6 October 1978. Using up and down time data from a nuclear power plant three dimensional plots were displayed which demonstrated the positive correlation and nature of the failure rate and repair rate. This research is still in process.

In analyzing the Bevalac Data it quickly became apparent that we should consider the effect of various shut-off policies. For example, a Computer Hardware failure will shut down the Magnet Power Supply, but not vice versa. As a consequence, we have recently completed work on a paper, "Availability Theory for Systems Whose Components are Subject to Various Shut-off Policies," [3]. In this paper we investigate in depth the consequences of various

shut-off policies. Availability and related formulas now involve more than the means of the distributions involved. However, comparisons with the case of all exponential life and repair distributions can be given in terms of IFR and DFR lifetime and repair distributions.

A post-doctoral student, Eduardo Subelman was partially supported by this grant in 1977-78. He has been working on the problem of determining the optimal maintenance policy for a system under the assumption of a finite planning horizon. He assumes that the expected operating and repair costs of a system are convex increasing functions of the time elapsed since the system was last maintained. Under these hypotheses he has determined the structure of policy that minimizes the expected operations and maintenance costs over the horizon. This policy has a particularly simple form, depending only on two critical numbers. The computation of these numbers can also be reduced to simple computations.

An invited paper on "Recent Developments in Reliability Theory," was presented at the *Seventh Conference on Stochastic Processes and Their Applications* at Twente University of Technology, Enschede, The Netherlands, 15-19 August 1977. An abstract of the talk is attached with this report.

Professor Kjell Doksum was partially supported by this grant during summer 1978. His paper on "Rank Tests for the Matched Pair Problem with Life Distributions" is partially credited to this grant.

Professor Sheldon Ross has credited his paper, "Approximations in Finite Capacity Multi-Server Queues with Poisson Arrivals," in part to this grant.

References

- [1] Barlow, R. E. and T. Liang, "Availability Analysis of the Superhilac Accelerator," ORC 77-21, Operations Research Center, University of California, Berkeley, (July 1977).
- [2] Chang, M. K., "Failure Data Analysis of the Superhilac Radio Frequency Subsystem," Operations Research Center Report, Operations Research Center, University of California, Berkeley, California.
- [3] Barlow, R. E. and E. Sid, "Availability Theory for Systems Whose Components are Subject to Various Shut-Off Policies," Operations Research Center, University of California, Berkeley, California.

REPORTS SUPPORTED BY THE ARMY RESEARCH OFFICE

- Barlow, R. E. and T. Liang, "Availability Analysis of the Superhilac Accelerator," ORC 77-21, Operations Research Center, University of California, Berkeley, (July 1977).
- Block, H. W., "Monotone Hazard and Failure Rates for Multivariate," ORC 75-20, Operations Research Center, University of California, Berkeley, (October 1975).
- Chang, M. K., "Failure Data Analysis of the Superhilac Radio Frequency Subsystem," Operations Research Center Report, Operations Research Center, University of California, Berkeley, California.
- Derman, C., G. J. Lieberman and S. M. Ross, "A Renewal Decision Problem," ORC 76-28, Operations Research Center, University of California, Berkeley, (September 1976).
- Doksum, K., "Rank Tests for the Matched Pair Problem with Life Distributions," Department of Statistics, University of California, Berkeley, California.
- Glassey, C. R. and S. M. Ross, "Queueing Models for Multiple Chamber Locks," ORC 76-27, Operations Research Center, University of California, Berkeley, (September 1976); Transportation science, Vol. 10, No. 4, pp. 391-404, (November 1976).
- Jewell, W. S., "Bayesian Life Testing Using the Total Q on Test," ORC 76-3, Operations Research Center, University of California, Berkeley, (January 1976); Proceedings: Conference on Theory and Application of Reliability, University of Southern Florida, Tampa, (December 15-18, 1975).
- Nozaki, S. A. and S. M. Ross, "Approximations in Multi-Server Poisson Queues," ORC 76-10, Operations Research Center, University of California, Berkeley, (April 1976).
- Nozaki, S. A. and S. M. Ross, "Approximations Finite Capacity Multi-Server Queues with Poisson Arrivals," O C 77-34, Operations Research Center, University of California, Berkeley, (December 1977).
- Smith, D. R., "Optimal Repairman Allocation Models," ORC 76-7, Operations Research Center, University of California, Berkeley, (March 1976).

PERSONNEL SUPPORTED

October 1, 1975 - September 30, 1978

Professor Richard E. Barlow

Professor Kjell Doksum

Professor Sheldon M. Ross

William M. Alexander

Mary F. Burk

Bernard Davis

Jaw Huan Hsiung

Robert D. Levin

Shun-Chen Niu

Shirley Nozaki

Zvi Schechner

Eduardo Subelman

Randall Willie

Alexander Wu

DEGREES AWARDED

October 1, 1975 - September 30, 1978

William M. Alexander, Ph.D.
December 1978*

Mary F. Burk, M.S.
June 1976

Robert D. Levin, Ph.D.
June 1977

Shun-Chen Niu, Ph.D.
December 1977

Eduardo Subelman, Ph.D.
December 1976

* Thesis was filed August 26, 1978 and William Alexander will receive his degree December 1978.

ABSTRACTS

AVAILABILITY ANALYSIS OF THE SUPERHILAC ACCELERATOR

Analysis of Two Years of Operating Data from the
Lawrence Berkeley Laboratory

by

Richard E. Barlow
Department of Industrial Engineering
and Operations Research
University of California, Berkeley

and

Tom Y. Liang
Operations Research Center
University of California, Berkeley

Records of 26 months of operating, repairing, and maintaining the nuclear accelerator Superhilac at the Lawrence Berkeley Laboratory (LBL) are analyzed with respect to system availability and reliability.

A major portion of the report is devoted to building a suitable model for the availability analysis. Some specific recommendations for improvement are also given. The current availabilities for operating the machine in Modes 1, 2, and 3 are 64.8%, 76.9%, and 80% respectively. The Adam injector is most responsible for causing the current low availability level of Mode 1. An increase of 15 hours on top of the current MTBF of 22.74 hours for the Adam injector (*or*, a decrease of about 2 hours from its current MTTR of 5.15 hours) would result in an overall boost of 5.6% for Mode 1 availability. (See Chapter 3 and Appendix 2 for additional recommendations.) Either way, this would translate into about 6 more usable hours every 4 days of operations for Mode 1.

Some optimization schemes (when the costs associated with improving the various equipments are known) are indicated for obtaining specific adjustments to be made systemwide in order to achieve economically an assigned higher availability level. Such models and schemes can easily be adapted to future needs as the information gathering system evolves in the near future into an organic computer network capable of collecting and analyzing much more voluminous and accurate data.

Also included are system reliability and miscellaneous findings from Time Series Analysis and Total Time on Test Plots.

AVAILABILITY THEORY FOR SYSTEMS
WHOSE COMPONENTS ARE SUBJECT
TO VARIOUS SHUT-OFF POLICIES

by

R. E. Barlow

E. Sid

Barlow and Proschan [1] derived limiting series system availability assuming non-field components are in "suspended animation" at system failure and until repair is completed. Limiting results in this case depend only on distribution means. When this model is modified so that some non-failed components continue to operate after system failure, limiting availability is no longer just a function of distribution means. Consider a two component series system with failure distributions. F_1 , F_2 and repair distributions G_1 , G_2 . If either F_1 or G_1 is IFR (DFR) and all other distributions are exponential, limiting availability is *greater (less)* than in the exponential case. If G_2 is IFR (DFR) and all other components are exponential, limiting availability is *less (greater)* than in the exponential case. For F_2 general and all other distributions exponential, limiting availability is the same as in the exponential case. Availability formulas for series systems with more than two components and various shut-off policies are obtained for exponential case.

Reference

- [1] Barlow, R. E. and F. Proschan, "Availability Theory for Multicomponent Systems," Multivariate Analysis III, Academic Press, pp. 319-335, (1973).

MONOTONE HAZARD AND FAILURE RATES FOR MULTIVARIATE DISTRIBUTIONS

H. W. Block

ABSTRACT

It is shown that the monotone multivariate hazard and failure rates of Harris and of Brindley and Thompson have no natural analog involving the multivariate failure rate function of Basu for absolutely continuous distributions. Quantities related to the multivariate failure rate function are used to define monotone hazard and failure rates. It is shown that these are equivalent to the monotone hazard and failure rates of Harris and of Brindley and Thompson. Based on these quantities, the loss of memory property of Marshall and Olkin is characterized.

FAILURE DATA ANALYSIS OF THE
SUPERHILAC RADIO FREQUENCY SUBSYSTEM

by

Mark K. Chang

This report is a follow-up of the study done by Liang [1], [2] in 1977 to investigate new techniques for analyzing SuperHILAC system availability. Recent and more accurate data are used and emphasis is on the Radio Frequency (RF) subsystem and its components. Time Series Analysis and Total Time on Test plots are the main tools used in the analysis. Recommendations for the improvement of RF availability, general SuperHILAC performance, and the data collecting process are given. The primary result suggests that the RF operating period should be extended.

C. Derman
G. J. Lieberman
S. M. Ross

ABSTRACT

A system must operate for t units of time. A certain component is essential for its operation and must be replaced, when it fails, with a new component. The class of spare components is grouped into n categories with components of the i th category costing a positive amount C_i and functioning for an exponential length of time with rate λ_i . The main problem of interest is, for a given t , to assign the initial component and subsequent replacements from among the n categories of spare components so as to minimize the expected cost of providing an operative component for t units of time.

In Section 1 we show that when there are an infinite number of spares of each category, the optimal policy has a simple structure. Namely, the time axis can be divided up into n intervals, some of which may be vacuous, such that when a replacement decision has to be made it is optimal to select a spare from the category having the i th largest value of λC whenever the remaining time falls into the i th closest interval to the origin. In Section 2 we consider the situation where $n = 2$ and there is only a single spare of one category and an infinite number of the other. In Section 3 we consider the case where there is only a finite number of spares for certain of the categories under the assumption that a rebate is allowed for the component in use at the end of the problem. In Section 4 we allude to a generalization of the model in Section 1 allowing for discounting or for the possibility that the system may randomly terminate before the t units of time expire. An optimal policy has the same simple structure as in Section 1.

RANK TESTS FOR THE MATCHED PAIR PROBLEM WITH LIFE DISTRIBUTIONS

by

K. Doksum

We consider matched pair experiments where the measurements (x_i, y_i) in each pair are time measurements such as failure times in a reliability study. A matched pair experiment is a simple block design where pairs are formed to block out nuisance factors. For example, matched pairs could be two brands for the front tires of a car (where differences between cars are then blocked out), a side of a house with the left and right halves painted with two makes of paints, etc.

If the shape of the failure time distributions are not known, it makes sense to consider rank test. We considered three classes of such tests:

(i) The class of signed rank tests based on the signed ranks of $\{y_i - x_i\}$. Here a test based on an exponential scores statistic turns out to have good power properties in reliability models.

(ii) A class of signed rank tests based on the signed ranks of $\{\log y_i - \log x_i\}$. The Wilcoxon signed rank test has good power properties in reliability models for this class. We show that each isotonic signed rank test has isotonic power with respect to tail ordering as defined in [1].

(iii) Invariance considerations leads to conditional tests based on the joint ranks of $\{x_i, y_i\}$. In this class, an exponential score (Savage) statistic is optimal (in the minimax sense) over the class of IFRA distributions. Isotonic conditional rank tests have isotonic power with respect to tail ordering.

- [1] Doksum, K.A. (1969). "Starshaped transformations and the power of rank tests", Ann. Math. Statist., pp. 1167-1176.

QUEUEING MODELS FOR MULTIPLE CHAMBER LOCKS

C. R. Glassey

and

S. M. Ross

ABSTRACT

Several models for predicting mean waiting times of river traffic at a multiple chamber lock were developed and tested. Mean waiting times predicted by the M/G/1 model differed significantly from observed times. Analysis of possible causes of failure of this model suggested a limited queue length M/G/1 model for one chamber, from which more accurate predictions were derived. For the two chamber system, an M/G/1 model with random batch size was developed. This model yields a lower bound for mean waiting time. These last two models can be used to predict system performance under various operating conditions.

BAYESIAN LIFE TESTING USING THE TOTAL Q ON TEST

W. S. Jewell

ABSTRACT

Suppose the basic shape of the cumulative failure (hazard) function has been identified for a certain component, and that an unknown parameter θ for a new production run of similar components is to be estimated. In particular, suppose that the failure function is of *proportional* type, $R(x) = \theta Q(x)$, where Q is the known shape function, and that θ is sampled from a prior gamma density. By using a new statistic, called the *total Q on test* (TQT), it is possible to perform Bayesian updating during a variety of lifetime testing programs in a manner similar to total time on test plots. This statistic can also be used with complete lifetime data, extending over several product runs, to identify the failure form Q , and to estimate the gamma hyperparameters. Extensions include the use of several TQT statistics to estimate the relative strength of competing hazard functions.

APPROXIMATIONS IN MULTI-SERVER POISSON QUEUES

Shirley A. Nozaki

and

Sheldon M. Ross

ABSTRACT

Our major objective is to obtain an approximation for the average time spent waiting in queue by a customer in an $M/G/k$ queueing system--call it W_Q . This is done by means of an approximation assumption presented in Section 2, which is shown to be asymptotically valid both in heavy and in light traffic. In Section 3, the approximation assumption is used to derive an approximation for W_Q . Numerical comparison with tables given by Hillier-Lo in the special case of Erlang service times indicate that the approximation, which depends on the service distribution only through its first two moments, works remarkably well. In addition, as a by-product of our analysis, we also obtain approximations for the distribution of the number of busy servers and the mean length and number of customers in a busy period. These latter approximations depend on the service distribution only through its mean.

In Section 4, we show that the approximation assumption is valid and leads to the exact result in the case of a limited capacity system where no queue is allowed to form.

APPROXIMATIONS IN FINITE CAPACITY
MULTI-SERVER QUEUES WITH POISSON ARRIVALS

Shirley A. Nozaki

and

Sheldon M. Ross

ABSTRACT

In this paper, we consider an $M/G/k$ queueing model having finite capacity N . That is, a model in which customers, arriving in accordance with a Poisson process having rate λ , enter the system if there are less than N others present when they arrive, and are then serviced by one of k servers, each of whom has service distribution G . Upon entering, a customer will either immediately enter service if at least one server is free or else join the queue if all servers are busy. Our results will be independent of the order of service of those waiting in queue as long as it is supposed that a server will never remain idle if customers are waiting. To facilitate the analysis, however, we will suppose a service discipline of "first come first to enter service."

OPTIMAL REPAIRMAN ALLOCATION MODELS

D. R. Smith

ABSTRACT

A system of n components under the care of one repairman is modeled. The components are subject to failure, whereupon they may be repaired one at a time. It is desired to repair failed components in such a manner that the ergodic probability that the system works is maximized.

It is assumed that each component and the system as a whole can be either working or failed, with the relationship between the working of the system and the working of the components given by a coherent structure function. The time a component works, or the time to repair a component is an exponential random variable of known rate. All components are independent, and at most one component may be under repair at a given time.

Although the general problem is in principle soluble by known methods, computational difficulties are enormous for moderate sized systems. In addition, such methods give no general insight into the structure of the optimal policy. Therefore, bounds and approximations for general systems are highly useful.

One bound for the optimal ergodic probability that the system works is given by the ergodic probability that the system works under a particular policy. The time reversible policy given yields easily obtainable ergodic probabilities for all states, and is useful for bounding purposes.

Most real systems are highly reliable in nature. Parametrization of the rates of the exponential random variables given earlier allows investigation of asymptotic system properties as the system becomes very reliable. Specifically, for a given policy, the asymptotic ergodic probability of all states and the asymptotic passage times between states may be computed. These results allow one to obtain the asymptotic optimal unreliability of an arbitrary system, and to obtain the asymptotically optimal policy for assignment of the repairman in many cases. Intuitively, the asymptotically optimal policy is close to optimal for highly reliable systems.

Although highly unreliable systems occur less frequently, such systems may be treated in a similar manner with similar results.

Two specific examples of systems are treated in the paper: the series system and an arbitrary system of stochastically identical components. The series system occurs often in practice, since many simple systems cannot tolerate failure of any components.

These two examples lead to relatively simple solutions. Arbitrary systems generally do not. The two component parallel system yields a fairly complicated criterion for choosing between the two potentially optimal policies.

When the series system is composed of components whose failure rates are identical, the ergodic probability that the system works is independent of policy.

For the two component series system, it is optimal to repair the longer expected lifetime component first, and this is true even if the repairman is subjected to random intervals during which he is not allowed to work.

For an n component series system, the optimal policy seems to be to repair the components in order of increasing expected lifetimes. This result can be proven if the optimal policy can be written as a list, but a more general proof seems to be elusive.

When a system is composed of stochastically identical components, it is often possible to eliminate most policies from consideration. Two examples of this technique are given, including one in which the optimal policy is explicitly obtained.

Seventh Conference on Stochastic Processes and their Applications

Twente University of Technology, Enschede,
The Netherlands, 15-19 August 1977

Reprinted from

ADVANCES IN APPLIED PROBABILITY

Vol. 10, No. 2, pp. 255-338

June 1978

I. INVITED REVIEW AND RESEARCH PAPERS

Recent developments in reliability theory

RICHARD. E. BARLOW, *University of California at Berkeley*

The key paper of Block and Savits (1976) has been used by Ross (1977) to extend the IFRA closure theorem (Birnbaum, Esary and Marshall (1966)) to multi-state coherent systems. The key result is that if T_1, T_2, \dots, T_n are independent IFRA, then

$$(1) \quad \left[E h^\alpha \left(\frac{T_1}{\alpha}, \frac{T_2}{\alpha}, \dots, \frac{T_n}{\alpha} \right) \right]^{1/\alpha}$$

is \downarrow in $\alpha > 0$ for all $h \geq 0$ and $h \uparrow$.

An immediate corollary is that if g is \uparrow and

$$\frac{1}{\alpha} g(\alpha T)$$

is \downarrow in $\alpha > 0$ for all $T = (T_1, T_2, \dots, T_n)$, then $g(T_1, T_2, \dots, T_n)$ is IFRA. For min cut-sets K_1, K_2, \dots, K_k and

$$g(T) = \min_{1 \leq s \leq k} \max_{i \in K_s} T_i$$

we have the IFRA closure theorem. For $g(T) = T_1 + T_2 + \dots + T_n$ we have the IFRA convolution theorem.

Block and Savits (1977) define T_1, T_2, \dots, T_n to be multivariate IFRA if (1) holds. Let $X_i(t)$ be the state $\{0, 1, 2, \dots, n\}$ of component i . If $X_i(t)$ is \downarrow in $t \geq 0$,

$$T_k = \sup \{t \mid X_i(t) \geq k\}$$

and T_1, T_2, \dots, T_m are multivariate IFRA, then $\{X_i(t); t \geq 0\}$ is an IFRA failure process.

If $\phi \uparrow$, then

$$\phi[X_1(t), X_2(t), \dots, X_n(t)]$$

is also an IFRA failure process.

Bo Bergman in Sweden (see e.g. Bergman (1978)) obtains optimal maintenance policies for multi-state components. Bob Smith at Columbia University obtains optimal policies for coherent systems. Barlow and Campo (1975) show how IFRA and other classes may be tentatively identified from data.